Model-based measurement of epileptic tissue excitability.

Paul Frogerais, Jean-Jacques Bellanger, Fabrice Wendling

Model-based measurement of epileptic tissue excitability

P. Frogerais, J.J. Bellanger, and F. Wendling

Model-based measurement of epileptic tissue excitability

P. Frogerais, J.J. Bellanger, and F. Wendling

***Abstract*— In the context of pre-surgical evaluation of epileptic patients, depth-EEG signals constitute a valuable source of information to characterize the spatiotemporal organization of paroxysmal interictal and ictal activities, prior to surgery. However, interpretation of these very complex data remains a formidable task. Indeed, interpretation is currently mostly qualitative and efforts are still to be produced in order to quantitatively assess pathophysiological information conveyed by signals. The proposed EEG model-based approach is a contribution to this effort. It introduces both a physiological parameter set which represents excitation and inhibition levels in recorded neuronal tissue and a methodology to estimate this set of parameters. It includes Sequential Monte Carlo nonlinear filtering to estimate hidden state trajectory from EEG and Particle Swarm Optimization to maximize a likelihood function deduced from Monte Carlo computations. Simulation results illustrate what it can be expected from this methodology.**

I. Introduction

During, pre-surgical examination of epileptic patients, diagnostic is mainly based on merging information from anatomo-functional imaging, semiology and from electrophysiological signals recorded from scalp-electrodes (EEG signals) or depth-electrodes (SEEG signals). The latter capture important information about dynamical electrical activities arising from neuronal populations close to electrode contacts (2 mm long, 0.8 mm diameter for intracerebral electrodes). Interpretation of recorded signals is a crucial issue that is addressed, in this paper through modeling. The goal is to relate various temporal patterns observed in depth-EEG signals during interictal/ictal to modifications of model parameter values. These parameters can be interpreted in the model as pathological modifications of excitation and inhibition efficiencies. In order to establish such a relationship, parameters must be estimated from real observations. After description of the model (section II), we present a new identification methodology (section III) which is essentially based on likelihood computations through Monte Carlo (MC) sequential Bayesian filtering. In section IV simulation results are given and discussed before conclusion.

II. Model description

The model we introduce here belongs to the class of lumped-parameter models [1] introduced in the 70s to describe background activity or evoked potential responses. Here, it was adapted to hippocampus activity in epilepsy [6]. In the cortical tissue, distinct neuronal subpopulations types can be distinguished. The interactions between these subpopulations are either inhibitory or excitatory. The electrical activity they develop can be modeled as illustrated on fig. 1. Three subpopulations of neurons are considered: excitatory pyramidal neurons (*Pe1* and *Pe2*) and two types of interneuron providing either slow dendritic inhibition (*Psi*) or fast somatic inhibition (*Pfi*). Coefficients C_i correspond to (mean) numbers of synaptic contacts between subpopulations and are supposed known and time invariant. Each subpopulation module includes one (or more) linear filtering operator(s) whose output(s) is (are) applied to a nonlinear no memory operator *S(.)*. This operator mimics threshold and saturation effects occurring at the soma. Input of filters and outputs of nonlinear operators represent the mean firing rate of action potentials. Outputs of filters represent excitatory or inhibitory (when weighted by -1) post synaptic membrane potentials resulting from time averaging in dendrites of impulse synaptic currents induced by afferent population(s). Biophysically, the EEG signal (field potential) recorded with an depth-EEG contact depends primarily on postsynaptic potential variations in *Pe1* and *Pe2* pyramidal cells. In such signal the influence of *Psi* and *Pfi* electrical activities can be neglected. Only three distinct transfer functions denoted *he*, *hsi* and *hfi* are introduced here. Their generic Laplace transfer function is

$$h(s)=\frac{1}{(\tau_k+s)^2} \quad \text{where} \quad \tau=\tau_e\,,\ \tau=\tau_{si} \quad \text{or} \quad \tau=\tau_{fi}<\tau_{si}$$

are time constants whose respective values are fixed in accordance with those reported in literature [6]. Each transfer function introduces two scalar state variables. The influence of cortical neighborhood random activity on the four local subpopulations is resumed by a positive-mean time continuous Gaussian white noise *W(t)* applied on *he* input of *Pe1* in addition to *Pe2* output activity. The three

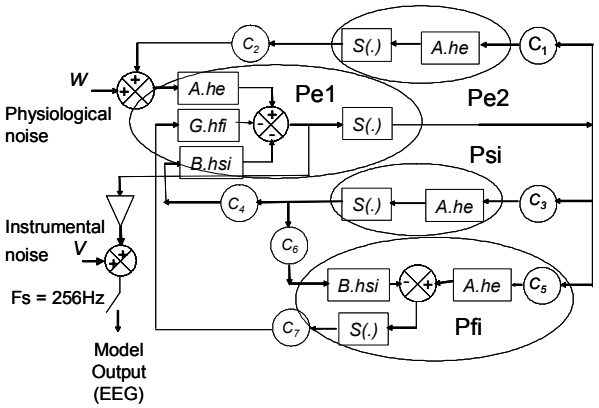

Fig. 1. The SEEG model.

Manuscript received April 16, 2007. P. Frogerais, J.J Bellanger and F. Wendling are with INSERM, U642, Rennes, F-35000, France; Université de Rennes 1, LTSI, Rennes, F-35000, France (corresponding author: +33 223 235 605; e-mail: paul.frogerais@univ-rennes1.fr).

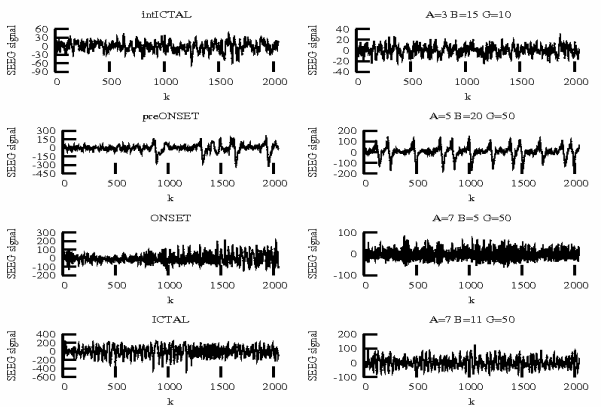

Fig. 2. (left) segmented signals recording during an epileptic seizure. (right) simulated signals by the model (Fig. 1) with different values of the parameters (A,B,G).

positive parameters A, B and G are interpreted as positive synaptic excitation gain, synaptic slow inhibition gain and synaptic fast inhibition gain, respectively. These quantities are those we want to measure as they are supposed to vary during the transition to seizures. Therefore, the 3D parameter $\theta=(A,B,G)\in\mathbb{R}^3$ must be estimated from real signals in order to 'observe' these modifications. Finally instrumentation high pass filter and additive observation Gaussian noise are included in the model before sampling operator and the overall system can be written:

$$dX = f(X,\theta)dt + G(\theta)d\beta_t \quad (1)$$

$$Y_k = HX_k + v_k \quad (2)$$

where (1) is a vector ($X\in\mathbb{R}^{14}$) stochastic differential equation (SDE) with a linear diffusion term $G(\theta)d\beta_t$ and a drift term $f(X,\theta)dt$. The Brownian motion increment $d\beta_t$ is the centered version of $W(t)dt$. Vector X regroups state variables of subpopulation input filters plus two variables for the output high-pass filter. The function f accounts for i) nonlinearities (sigmoid S(.) functions), ii) transfer functions and iii) parameter θ. The output Y_k is a linear form of sampled state vector X_k. In fig. 2 four real depth EEG segments recorded before and during onset of a seizure (left) may be visually compared with model simulations (right) obtained with empirically adjusted values of (A,B,G).

III. Identification Method

A. Discrete scheme for the state equation.

To simulate time continuous SDE system (1) with discrete time observation (2), the SDE can be discretized by a second order Runge-Kutta method. In our particular model, we showed that this discrete approximation leads to:

$$X_k = f_{rk2}(X_{k-1},\theta,\Delta) + G_{rk2}(\Delta,\theta)W_k \quad W_k \sim N(0,\sigma_k) \quad (3)$$

The two identification methods presented below and illustrated on fig. 3 use discretization (3) and equation (2).

B. First method: estimated moment method.

The first method involves only boxes b) and c) in figure

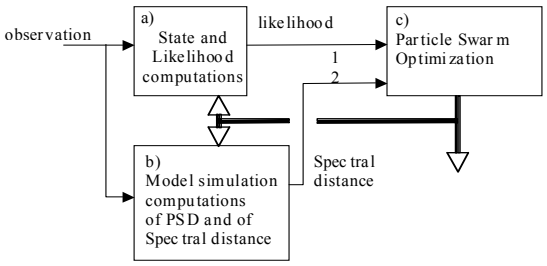

Fig. 3. The identification method.

3). It consists in comparing estimated values $\hat{\theta}$ with the observed sampled signal $y_{1:N}$ (modeled as an outcome of the random vector $Y_{1:N}=[Y_1,\ldots,Y_N]$) through a feature $F(\theta)$ define as a functional of the probability law $P^{\theta}_{Y_{1:N}}$. According to the moment parameter estimation method [8] an estimation $\hat{\theta}$ can then be defined as $\hat{\theta}=\arg\min_\theta \left\|\hat{F}(y_{1:N})-F(\theta)\right\|$ where $\hat{F}(Y_{1:N})$ is an estimation of $F(\theta)$ computed on observation and $\|.\|$ a quadratic norm. Here we retain $F=[F_1,F_2,F_3]$ where the $F_i,i=1,2,3$ are normalized expected powers of four filtered versions of observation $Y_{1:N}$ obtained by band-pass filtering through three frequency bands: delta (0-4Hz), theta and alpha (4-12Hz), beta and gamma (12-64Hz). As $F(\theta)$ can not be calculated analytically, we estimate it as function $\hat{F}_S(y_{S1:n_S}(\theta,W))$ of an outcome of model output (2) (n_S time samples) simulated with parameter value θ and input noise sequence W in (3). Because the function $\theta\rightarrow\hat{\theta}=\arg\min_\theta \left\|\hat{F}(y_{1:N})-F(\theta)\right\|$ may have several local minima a Particle Swarm Optimization Algorithm (PSOA) [7] is then utilized to compute $\hat{\theta}$ (box c) fig.3).

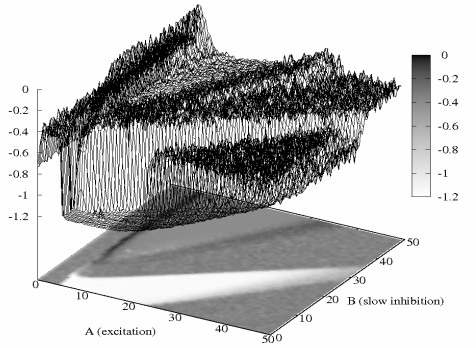

Fig. 4. The function $-h(\theta,w_k)$ calculated in the A-B plane ($G=50$) on a realization of the second model ($A=5$, $B=20$,$G=50$).

A PSOA is a global optimization procedure that propagates a set of K candidate θ values $\theta^k_j, k=1..K$ during iterations $j=1..J$ before stopping. At iteration j and for each particle k, its *performance* is valuated by computing

$h(\theta_j^k, W) = \left\| \hat{F}(y_{1:N}) - \hat{F}_S(y_{S1:n_r}(\theta_j^k, W) \right\|$ which is random since it depends on noise outcome W (Fig. 4 illustrates the effect of this randomness which entails difficulties to find an reliable optimum). For each j value the PSOA compute a small displacement for each particle as a function of its current position, of its position at preceding iteration and of best encountered positions in the past for itself and for a currently and randomly defined set of some other particles in the swarm (named advisors). The algorithm hence combines different randomly selected candidate values to produce new values in the same manner as evolutionary algorithms. It is stopped when h values became stable.

C. *Second method: Maximum likelihood method.*

This method involves mainly boxes a) and c) in fig.3). It necessitates an estimation $\hat{L}(y_{1:N}, \theta)$ of the likelihood $L(y_{1:N}, \theta)$ which cannot be computed analytically. This approach was proposed in [1] for neural mass models simpler than our model (only three subpopulations) and in a non-pathological context: it consists in implementing a nonlinear Bayesian filter for state point estimation (a form of improved extended Kalman filter [9]) and to compute likelihood as a function of innovations (prediction errors of the observation) for each θ value. Here we propose to utilize a particle filter [2] which provide information concerning state conditional probability distribution given observations. Hence we compute the likelihood by mean of a particle filter included in a) fig.3) and values of θ that lead to large likelihood values are, as for the first method, obtained with a PSOA algorithm (c) fig. 3)). In order to limit the particle filter computational time, the parameter space research is reduced by forbidding θ values (action of b) on c) fig.3)) which do not respect the constraint $-h(\theta,W) > \alpha$ where α is set to an empirical value ($\alpha = -0.2$). This constraint is faster to calculate than the particle filter.

Non-linear Bayesian filtering methods [9] are used to estimate the hidden state trajectory $X(t)$ of a non-linear Markov system observed through a noisy no memory function $y(t) = g(X(t), noise)$ of this state. In case of discrete time state evolution, these methods consist in estimating the probability density $p_{X_k|Y_{1:k}}$ of the state X_k at discrete time t_k given discrete observations $Y_{1:k}$. Generally it is a difficult problem, which necessitate approximate numerical resolutions, as doing by particular filters. Particular filtering [3] is an application of a MC method called sequential importance sampling with re-sampling (SISR) algorithm. This recent and popular method sample at each time k a set $\hat{x}_k^i$, $i=1,...,N_s$, from an auxiliary density, the instrumental density q, and compute associated weights w_k^i. These weights may be chosen such that the expectation $E[f(X_k)|Y_{1:k}]$ can be approximated by $\hat{E}[f(X_k)|Y_{1:k}] = \sum_{i=1}^{N_s} \omega_k^i f(x_k^i)$ for arbitrary $f(.)$ when N_s is large [2]. Hence, when N_s is large the information provided by the set $(\hat{x}_k^i, w_k^i)$, $i=1,...N_s$, is equivalent to the $p_{X_k|Y_{1:k}}$ knowledge. A natural and classical choice for the instrumental law is $q = p_{X_k|X_{k-1}}$ which correspond to the so-called bootstrap filter [3]. But for reasonable (not too large) N_s value degeneracy phenomena appear: after several steps, only a very little subset of particles give valuable information on $p_{X_k|Y_{1:k}}$. So, we used here a more sophisticated importance law called optimal instrumental density [2]. Particle filtering with optimal importance density was applied to discretized version (3) of system (1).

Furthermore it can be shown [4] that the likelihood can be approximated by:

$$\hat{L}(y_{1:N}, \theta_1) = \sum_{k=1}^{N} \log \sum_{i=1}^{N_s} w_k^i \qquad (4)$$

To illustrate, fig. 5 shows estimated values of the log-likelihood as a function of (A,B) and with $N_s=20$ on a signal simulated with $\theta = (5,20,50)$. G was set to its real value ($G=50$). We can notice several local maxima and the global maximum argument (the ML estimation) proximal to the real value $(A=5,B=20)$.

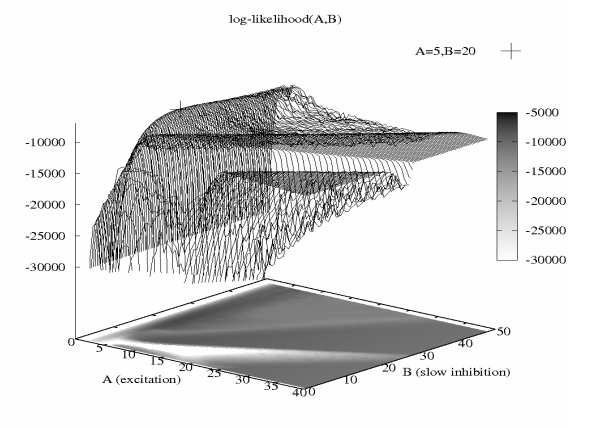

Fig. 5. The log-likelihood calculated on a realization of the first model $A=5, B=20, G=50$.

IV. Results on simulated data

We focused on four different models corresponding to four different activities shown in fig. 2 (right). For each model, ten realizations were simulated. Then the second parameter estimation method, presented in the previous section, was applied on these 40=10x4 signals, with $N_s=20$. A 40 particles swarm with 3 advisors for each particle was used in the PSOA. Estimated parameter values are plotted in the parameter space, fig. 6. Different symbols are used to mark different models. Each set of links represents ten estimations obtained with ten output outcomes of a given model. We can note that, despite the estimation dispersion,

four sets can be easily distinguished. This shows the relevance of the estimation procedure.

TABLE I

SIGNALS		ML ESTIMATION		SENSITIVITY
		Mean	σ	$S_i(\theta)$
intICTAL	A	3.47	0.267	$S_1(\theta_1)$= 8.64e-3
θ_1 =(3,15,10)	B	25.2	9.87	$S_2(\theta_1)$= 1.003e-5
	G	33.7	17.9	$S_3(\theta_1)$= 2.18e-6
preONSET	A	4.9	0.062	$S_1(\theta_2)$= 57.9
θ_2 =(5,20,50)	B	20	0.325	$S_2(\theta_2)$= 1.08
	G	48.2	1.85	$S_3(\theta_2)$= 0.816
ONSET	A	7.43	0.414	$S_1(\theta_3)$= 3.1
θ_3 =(7,5,50)	B	5.24	0.323	$S_2(\theta_3)$= 10.06
	G	53.3	16.4	$S_3(\theta_3)$= 0.176
ICTAL	A	7.05	0.229	$S_1(\theta_4)$= 0.525
θ_4 =(7,11,50)	B	11	0.248	$S_2(\theta_4)$= 0.437
	G	45.3	17.7	$S_3(\theta_4)$= 1.101e-3

Parameter estimation on ten simulated signals for different activities (models) (intICTAL, preONSET, ONSET, ICTAL).

The experimental means and standard deviations of the four sets of parameter estimations are reported in table I. Except for the second model, the estimation of parameter G presents high variations. In the first model, B and G estimates have also a high variance.

In order to compare the two identification methods described in section III, parameters were estimated ten times for each model and for each of the two methods. For each model, this process was performed on the same simulated observation. The obtained experimental means and standard deviations are reported in table II. Globally the first estimator shows more dispersion than the second one. This dispersion is due to the dependency on W of the feature vector estimate. It could be reduced by taking n_S larger than observation duration but it can be shown that this would also introduce bias.

TABLE II

SIGNALS		ML METHOD		MOMENT METHOD	
		Mean	σ	Mean	σ
intICTAL	A	3.42	0.338	16.26	9.24
(3,15,10)	B	29.96	13.2	15.9	7.71
	G	53.6	20.5	38	31.6
preONSET	A	4.92	0.0362	5.72	0.375
(5,20,50)	B	20.8	0.143	24.2	2.36
	G	50.5	0.995	64.8	12.6
ONSET	A	7.38	0.713	23.9	6.71
(7,5,50)	B	5.06	0.663	6.72	2.03
	G	50.8	27	48.8	12.3
ICTAL	A	7.2	0.436	11,1	24,3
(7,11,50)	B	11.4	0.430	24,3	29,9
	G	34.4	18.2	29.9	8.2

Comparison of two parameter estimation methods computed 10 times on the same simulated signal. The real parameter values are in brackets.

For the second method, the MC sampling is the essential cause of dispersion. It can be reduced by increasing N_s and, therefore, proportionally increasing the computational time.

In order to numerically evaluate how a small variation $\Delta\theta$ on parameters has an effect on the EEG output, we simulated, with two models: $M_1=M(\theta)$ and $M_2=M(\theta+\Delta\theta)$ and for a same realization W of input noise, two signals $y_{s1:ns}(\theta W)$ and $y_{s1:ns}(\theta+\Delta\theta,W)$. An measure of the sensitivity is obtained by calculating the mean square error between these signals:

$$S_i(\theta) = \left\| y_{s1:ns}(\theta, n_r) - y_{s1:ns}(\theta + \Delta\theta_i, n_r) \right\|^2 \quad (5)$$

for $i=1,2,3$ with $\Delta\theta_1=(0.2,0,0)$, $\Delta\theta_2=(0,0.2,0)$, $\Delta\theta_3=(0,0,0.2)$. An evaluation of this quantity for a realization of each simulated model (fig. 2) is reported in Table I. Note that identifiability of the model increases when this parameter increases.

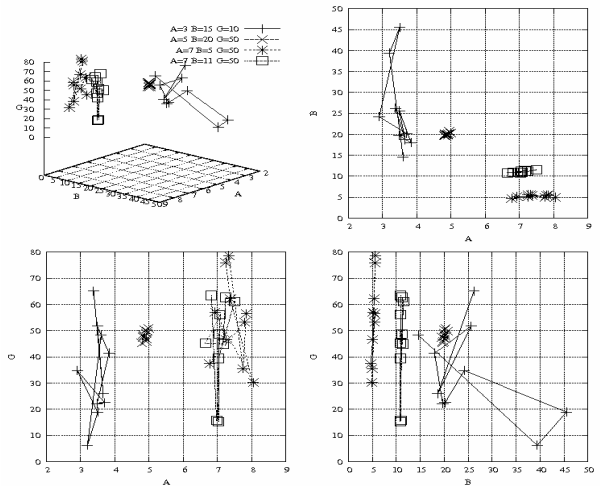

Fig. 6. Results of parameter estimation with simulated data. Ten signals were simulated by the model for 4 different values of theta corresponding to the different cases on the Fig. 1.

V. CONCLUSION

Results obtained on simulated signals show that estimation of synaptic gains is not easily achieved in some regions of the parameter space. Nevertheless, the ML method with MC approximation and particle swarm optimization we presented here makes this estimation feasible. The main difficulty before being able to apply it on larger databases is to address the problem of required computer time, more especially in case where model outputs are less sensitive to parameter vector values.

REFERENCES

[1] P.A. Valdes, J.C. Jimenez, J. Riera, R. Biscay, T. Ozaki "Nonlinear EEG analysis based on a neural mass model" Biol. Cybern. 81, 415-424 (1999)

[2] Olivier Cappé, Eric Moulines, and Tobias Ryden. Inference in Hidden Markov Models (Springer Series in Statistics). Springer-Verlag New York, Inc., Secaucus, NJ, USA, spinger edition, 2005.

[3] N.J. Gordon, D.J. Salmond, and A.F.M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE Proceedings, 1993.

[4] M. Hürzeler and H. R. Künsch. Approximating and maximizing the likelihood for a general state-space model. 2001.

[5] Joshua Wilkie. Numerical methods for stochastic differential equations. Physical Review E (Statistical, Nonlinear, and Soft Matter

[6] F. Wendling, A. Hernadez, J.J. Bellanger, P. Chauvel, F. Bartolomei "Interictal to ictal transition in human temporal lobe epilepsy: Insights from a computational model of intracerebral EEG" JCN, 2005.

[7] M. Clerc. Particle swarm optimization. ISTE, 2006.

[8] B Porat., Digital processing of random signals, Prentice-Hall, 1994.

[9] A. H. Jazwinski, Stochastic processes and filtering theory, Academic Press, 1970.